

Two-Variable Ehrenfeucht-Fraïssé Games over Omega-Terms

Manfred Kufleitner

Jan Philipp Wächter

University of Stuttgart, FMI

Abstract. Fragments of first-order logic over words can often be characterized in terms of finite monoids, and identities of omega-terms are an effective mechanism for specifying classes of monoids. Huschenbett and the first author have shown how to use infinite Ehrenfeucht-Fraïssé games on linear orders for showing that some given fragment satisfies an identity of omega-terms (STACS 2014). After revisiting this result, we show that for two-variable logic one can use simpler linear orders.

1 Introduction

As many fragments of first-order logic correspond to sets of finite monoids, the study of such fragments often results in the study of algebraic properties of these monoids. A recurring problem in this connection is the question whether a given identity of omega-terms $u = v$ holds. The omega-terms u and v in the identity are finite words over an alphabet of variables with an additional ω -power. This ω -power is usually interpreted as a mapping of monoid elements which maps the element m of a finite monoid M onto its generated idempotent m^ω , i.e. the element $e = m^k \in M$ for the smallest positive number k such that $e = e^2$, which is known to exist for any element in a finite monoid. With this interpretation the identity $u = v$ holds in M if u and v yield the same element for any assignment of the variables occurring in u and v onto elements of M and later interpretation of the concatenation with M 's binary operation.

To avoid confusion with the second historical meaning of ω to denote the linear order type of all natural numbers, omega-terms will be called π -terms in this paper. This notation was introduced by Perrin and Pin [3] in the context of infinite words. Of course, the ω -power will also be called π -power instead.

Often one can use Ehrenfeucht-Fraïssé games in order to show that a certain identity holds in all monoids belonging to a certain first-order logic fragment. This is usually done by describing a winning strategy for Duplicator in the n -round Ehrenfeucht-Fraïssé game for an arbitrary n . This approach usually involves a certain amount of book-keeping to keep track of the remaining rounds. M. Huschenbett and the first author presented a

way to avoid this book-keeping by playing infinite long Ehrenfeucht-Fraïssé games [1]. Here, the π -power gets replaced by a certain linear order type, namely $\varrho = \omega + \zeta \cdot \eta + \omega^*$, where ω denotes the linear order type of the natural number, ω^* its dual, ζ that of all integers and η that belonging to the rational numbers. This leads to the notion of generalized words which are not necessarily finite anymore.

In this paper we are first going to revisit this result and then have a further look on how the linear ordering can be simplified when we restrict our consideration onto fragments of first-order logic with only two variables.

2 Preliminaries

Generalized Words. A central notion for this paper is that of generalized words. We therefore fix an alphabet Λ . A *generalized word* w over Λ is a triple $w = (P_w, \leq_w, l_w)$ where P_w is a set of *positions*, \leq_w is a linear order relation on P_w (i. e. a total, reflexive, transitive and anti-symmetric binary relation), and $l_w : P_w \rightarrow \Lambda$ is a labeling map. The set P_w of positions of w is called the *domain* of w , denoted by $\text{dom}(w) = P_w$. To simplify notation, we write $w(p)$ instead of $l_w(p)$ for a position $p \in P_w$. The *order type* of w is the isomorphism¹ type of the linear ordered set (P_w, \leq_w) . Any finite word $u = a_1 a_2 \dots a_n$ of length $n \in \mathbb{N}$ in the set of finite words over Λ , denoted by Λ^* , can be regarded as a generalized word: its domain is the set $\text{dom}(u) = \{1, 2, \dots, n\}$, the order relation is the natural order, and the labeling map labels position $i \in \text{dom}(u)$ by a_i . We therefore see finite words as generalized words in the following, and refer by “word” to generalized words instead of finite words. We follow the custom of identifying the order type of any finite word of length $n \in \mathbb{N}$ (i. e. the order isomorphism type of the naturally ordered set $\{1, 2, \dots, n\}$) with n .

We consider two (generalized) words u and v *isomorphic*, if there is an isomorphism $\iota : \text{dom}(u) \rightarrow \text{dom}(v)$ of linear ordered sets which is compatible with the labeling, i. e. for which $u(p) = v(\iota(p))$ holds for all $p \in \text{dom}(u)$. We do not distinguish between isomorphic words anymore, and denote the set of (isomorphism classes of) *countable* words by Λ^{LO} ; the LO is for *linear order*. As mentioned above, we consider Λ^* as a subset of Λ^{LO} .

Given a linear ordered set (T, \leq_T) of isomorphism type τ we can create the τ -power w^τ of any word $w \in \Lambda^{\text{LO}}$: the domain of w^τ is $\text{dom}(w) \times T$, the linear order relation \leq_{w^τ} is defined by $(p, t) \leq_{w^\tau} (p', t')$ if $t <_T t'$ (i. e. $t \leq_t t'$ and $t \neq t'$) or if $t = t'$ and $p \leq_w p'$, and the labeling is given by $w^\tau(p, t) = w(p)$. One should observe that this coincides with the usual n -power for a natural number $n \in \mathbb{N}$ over finite words.

Logic over Words. Next, we give a brief overview of logic over words. We fix a countably infinite $\mathbb{V} = \{x, y, z, \dots\}$ of first-order variables. A first-order logical formula over words is syntactically defined by

$$\varphi ::= \top \mid \perp \mid x = y \mid \lambda(x) = a \mid x < y \mid x \leq y \mid \neg \varphi \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \exists x \varphi \mid \forall x \varphi,$$

¹Two ordered sets are *isomorphic* if there is an order-preserving bijection between them.

with $x, y \in \mathbb{V}$ and $a \in \Lambda$. The set of all these formulae is **FO**. A *sentence* is a formula $\varphi \in \mathbf{FO}$ with $\text{FV}(\varphi) = \emptyset$ where $\text{FV}(\varphi)$ denotes the set of free variables in φ which are defined as usual.

For the semantics of first-order formulae we define an \mathbb{X} -*valuation* for a finite subset $\mathbb{X} \subseteq \mathbb{V}$ on a word $w \in \Lambda^{\text{LO}}$ as the pair $\langle w, \alpha \rangle$ where $\alpha : \mathbb{X} \rightarrow \text{dom}(w)$ is an arbitrary map. Whether the \mathbb{X} -valuation satisfies a formula $\varphi \in \mathbf{FO}$ with $\text{FV}(\varphi) \subseteq \mathbb{X}$ depends on the following interpretations:

- Variables get values in $\text{dom}(w)$; free variables are interpreted according to α .
- \top is always satisfied, and \perp is never satisfied.
- The logical operations \neg , \vee , and \wedge are interpreted as usual, just like the quantifiers \exists and \forall .
- The binary operators of comparison $=$, $<$ and \leq are interpreted according to the ordering of $\text{dom}(w)$ with respect to w 's order type.
- λ gets interpreted by the labeling map l_w of w .

If $\langle w, \alpha \rangle$ satisfies a formula $\varphi \in \mathbf{FO}$ it is called a *model* and we write $\langle w, \alpha \rangle \models \varphi$. Any word $w \in \Lambda^{\text{LO}}$ is identified by $\langle w, \emptyset \rangle$, which is unique. This allows for the notation $w \models \varphi$ for a first-order sentence φ .

Fragments. The first author and A. Lauser proposed a formal definition of logical fragments as sets of formulae satisfying certain closure properties [2]. For this paper (just like in [1]) we are going to use a slightly modified version of this idea. A first-order formula μ in which a special additional constant predicate \circ appears exactly once is called a *context*. The predicate \circ is intended as a placeholder which can be replaced by other first-order formulae. The result of replacing \circ in a context μ by a formula $\varphi \in \mathbf{FO}$ is denoted by $\mu(\varphi)$. The notion of a context now allows for the formal definition of a fragment.

Definition 2.1. A *fragment* is a non-empty set of formulae $\mathcal{F} \subseteq \mathbf{FO}$ such that for all contexts μ , formulae $\varphi, \psi \in \mathbf{FO}$, $a \in \Lambda$ and $x, y \in \mathbb{V}$ the following conditions are satisfied:

1. If $\mu(\varphi) \in \mathcal{F}$, then $\mu(\top) \in \mathcal{F}$, $\mu(\perp) \in \mathcal{F}$, and $\mu(\lambda(x) = a) \in \mathcal{F}$.
2. $\mu(\varphi \vee \psi) \in \mathcal{F}$ if, and only if, $\mu(\varphi) \in \mathcal{F}$ and $\mu(\psi) \in \mathcal{F}$.
3. $\mu(\varphi \wedge \psi) \in \mathcal{F}$ if, and only if, $\mu(\varphi) \in \mathcal{F}$ and $\mu(\psi) \in \mathcal{F}$.
4. If $\mu(\neg \neg \varphi) \in \mathcal{F}$, then $\mu(\varphi) \in \mathcal{F}$.
5. If $\mu(\exists x \varphi) \in \mathcal{F}$ and $x \notin \text{FV}(\varphi)$, then $\mu(\varphi) \in \mathcal{F}$.
6. If $\mu(\forall x \varphi) \in \mathcal{F}$ and $x \notin \text{FV}(\varphi)$, then $\mu(\varphi) \in \mathcal{F}$.
7. $\mu(x < y) \in \mathcal{F}$ if, and only if, $\mu(x \leq y) \in \mathcal{F}$.

These closure properties are satisfied by many classes of formulae which arise naturally in the study of first-order logic, such as **FO** itself, the set of first-order formulae with limited quantifier alternations (e. g. Σ_n and Π_n), and the set of those whose number of variables is bounded by a natural number m (denoted by \mathbf{FO}^m).

We define the *quantifier depth* $\text{qd}(\varphi)$ of a formula $\varphi \in \mathbf{FO}$ in the usual manner and say that a fragment \mathcal{F} has *bounded quantifier depth* if there is an $n \in \mathbb{N}$ such that

$\text{qd}(\varphi) \leq n$ for all $\varphi \in \mathcal{F}$. For any fragment \mathcal{F} and any natural number n we define $\mathcal{F}_n = \{\varphi \mid \varphi \in \mathcal{F}, \text{qd}(\varphi) \leq n\}$ as the corresponding fragment with bounded quantifier depth n . One should verify that \mathcal{F}_n indeed is a fragment. An example of such a fragment is FO_n , the set of first-order formulae whose quantifier depth is bounded by n .

3 Ehrenfeucht-Fraïssé Games

In this section we will have a closer look on Ehrenfeucht-Fraïssé games. These games can often be used to show whether a given fragment of first-order logic can distinguish two words, i. e. whether there is a formula in the fragment for which one word is a model while the other is not.

We define the set of quantifiers as $\mathcal{Q} = \{\exists, \forall, \neg\exists, \neg\forall\}$. Note, that we also consider the “negated quantifiers” $\neq \exists$ and $\neq \forall$ as quantifiers. Let \mathcal{F} be a fragment, $x \in \mathbb{V}$ and $Q \in \mathcal{Q}$, then the *reduct* of \mathcal{F} by Qx is the set

$$\mathcal{F}/Qx = \{\varphi \in \text{FO} \mid Qx\varphi \in \mathcal{F}\}.$$

Note that a reduct of a fragment is either empty or a fragment itself.

We call the Ehrenfeucht-Fraïssé game which is played over the fragment \mathcal{F} and on the words $u, v \in \Lambda^{\text{LO}}$ the \mathcal{F} -game on (u, v) . Configurations $(\mathcal{G}, \langle u, \alpha \rangle, \langle v, \beta \rangle)$ of such a game consist of

1. an iterated, non-empty reduct \mathcal{G} of \mathcal{F} ,
2. a \mathbb{X} -valuation $\langle u, \alpha \rangle$, and
3. a \mathbb{X} -valuation $\langle v, \beta \rangle$

for a *finite* subset $\mathbb{X} \subseteq \mathbb{V}$ of variables. Such a configuration is also called an \mathbb{X} -*configuration* of the game.

In the beginning the game is in the \emptyset -configuration (\mathcal{F}, u, v) . The game can be played any number – even an infinite number – of rounds. In each round a variable gets assigned a new value; this variable can either be a variable which so far had no value or a variable which already has one. In more detail: suppose the game is in the \mathbb{X} -configuration $S = (\mathcal{G}, \langle u, \alpha \rangle, \langle v, \beta \rangle)$. One round consists of two steps:

1. The first player, called *Spoiler*, chooses
 - a quantifier $Q \in \mathcal{Q}$,
 - a variable $x \in \mathbb{V}$ such that $\mathcal{G}/Qx \neq \emptyset$, and
 - a *quest* q .

The quest is either a position in $\text{dom}(u)$ if $Q \in \{\exists, \neg\forall\}$ or a position in $\text{dom}(v)$ if $Q \in \{\forall, \neg\exists\}$.

2. The second player, called *Duplicator*, replies with a position r in the domain of the other word (with respect to the quest).

The follow-up configuration $S[Qx, q, r]$ of the game is defined in Table 1. Note that the chosen variable x can, but does not need to be from \mathbb{X} .

	Spoiler	Duplicator	resulting configuration $S[\mathbf{Q}x, q, r]$
$\mathbf{Q}x = \exists x$	$q \in \text{dom}(u)$	$r \in \text{dom}(v)$	$(\mathcal{G}/\exists x, \langle u, \alpha[x/q] \rangle, \langle v, \beta[x/r] \rangle)$
$\mathbf{Q}x = \forall x$	$q \in \text{dom}(v)$	$r \in \text{dom}(u)$	$(\mathcal{G}/\forall x, \langle u, \alpha[x/r] \rangle, \langle v, \beta[x/q] \rangle)$
$\mathbf{Q}x = \neg \exists x$	$q \in \text{dom}(v)$	$r \in \text{dom}(u)$	$(\mathcal{G}/\neg \exists x, \langle v, \beta[x/q] \rangle, \langle u, \alpha[x/r] \rangle)$
$\mathbf{Q}x = \neg \forall x$	$q \in \text{dom}(u)$	$r \in \text{dom}(v)$	$(\mathcal{G}/\neg \forall x, \langle v, \beta[x/r] \rangle, \langle u, \alpha[x/q] \rangle)$

Table 1: A single round of the \mathcal{F} -game in configuration $S = (\mathcal{G}, \langle u, \alpha \rangle, \langle v, \beta \rangle)$.

The winning conditions for the two players are as follows:

- If Spoiler cannot choose a quantifier and a variable such that the corresponding reduct is not empty (i. e. \mathcal{G} does not contain any more quantified formulae), then Duplicator wins.
- If either player cannot choose a position because its word's domain is empty, the other player wins.
- If the game reaches an \mathbb{X} -configuration $(\mathcal{G}, \langle u, \alpha \rangle, \langle v, \beta \rangle)$ such that there is a literal (i. e. an unquantified formula) $\varphi \in \mathcal{G}$ with $\text{FV}(\varphi) \subseteq \mathbb{X}$ and $\langle u, \alpha \rangle \models \varphi$ but $\langle v, \beta \rangle \not\models \varphi$, then Spoiler wins.
- Duplicator wins, if Spoiler does not win. In particular Duplicators wins every game which goes on forever.

Whenever a player has won, the game stops immediately. Given the winning conditions either Spoiler or Duplicator has a winning strategy on the words (u, v) in the \mathcal{F} -game, the game is *determined*. For a configuration $S = (\mathcal{F}, \langle u, \alpha \rangle, \langle v, \beta \rangle)$ we write $\langle u, \alpha \rangle \lesssim_{\mathcal{F}} \langle v, \beta \rangle$ if Duplicator has a winning strategy on S , or $\langle u, \alpha \rangle \not\lesssim_{\mathcal{F}} \langle v, \beta \rangle$ if Spoiler has one. One should note that $\lesssim_{\mathcal{F}}$ is a preorder, i. e. a reflexive and transitive binary relation.

Suppose that the quantifier depth of all formulae in a fragment \mathcal{F} is bound by $n \in \mathbb{N}$. Because we lose one quantifier level in very round of the game, the \mathcal{F} -game can last at most n rounds. This means that the \mathcal{F}_n -game can be seen as an n -round version of the \mathcal{F} -game for *any* fragment. We are going to have a closer look on this kind of games. The following result from [1] is an adaption of the Ehrenfeucht-Fraïssé Theorem for fragments of bounded quantifier depth.

Theorem 3.1. *Let \mathcal{F} be a fragment of bounded quantifier depth. For all words $u, v \in \Lambda^{\text{LO}}$ the following are equivalent:*

1. $u \models \varphi$ implies $v \models \varphi$ for all sentences $\varphi \in \mathcal{F}$ and
2. Duplicator has a winning strategy in the \mathcal{F} -game on (u, v) .

The proof for the implication “2. \Rightarrow 1.” does not depend on the bounded quantifier depth of the fragment. The implication thus holds also for the unbounded case. For the implication “1. \Rightarrow 2.” the boundedness, however, is substantial. An example for this is an Ehrenfeucht-Fraïssé game on $(a^\zeta, a^{\zeta+\zeta})$ for an arbitrary $n \in \mathbb{N}$ where ζ denotes the natural order type of all integers. Here, Duplicator has a winning strategy for all

FO_n -games with $n \in \mathbb{N}$ which yields that $a^\zeta \models \varphi$ implies $a^{\zeta+\zeta} \models \varphi$ for all sentences $\varphi \in \text{FO}$. On the other hand, Spoiler has a winning strategy in the (infinite) FO -game.

To gain a version of the theorem which does not limit the quantifier depth of the fragment, one has to restrict the words u and v to a special subset of words. For this we define:

Definition 3.2. Let τ be a linear ordering. A word $u \in \Lambda^{\text{LO}}$ is τ -rational if it can be constructed from the finite words in Λ^{LO} using the operations of concatenation and τ -power, only.

Next we need to define some special order types:

Definition 3.3. The order types of the sets \mathbb{N} , \mathbb{Z} , \mathbb{Q} and $\mathbb{Z}_{\leq 0} = -\mathbb{N}$ ordered naturally are denoted by ω , ζ , η and ω^* . Let $a \in \Lambda$. The order type of the word $a^\omega(a^\zeta)^\eta a^{\omega^*}$ is $\varrho = \omega + \zeta \cdot \eta + \omega^*$, and $\sigma = \omega + \zeta + \omega^*$ is the order type of $a^\omega a^\zeta a^{\omega^*}$.

With these definitions it is now possible to state the following theorem from [1].

Theorem 3.4. Let \mathcal{F} be fragment. For all ϱ -rational words $u, v \in \Lambda^{\text{LO}}$ the following are equivalent:

1. $u \models \varphi$ implies $v \models \varphi$ for all sentences $\varphi \in \mathcal{F}$ and
2. Duplicator has a winning strategy in the \mathcal{F} -game on (u, v) .

The implication “2. \Rightarrow 1.” is a special case of the according implication for Theorem 3.1 which holds – as already mentioned – even if one drops the requirement of \mathcal{F} being of bounded quantifier depth. Therefore, a proof for this direction can be obtained along the lines of a proof for the classic version of the Ehrenfeucht-Fraïssé Theorem. At the heart of the proof for the implication “1. \Rightarrow 2.” lies the following lemma:

Lemma 3.5. Let \mathcal{F} be a fragment, $x \in \mathbb{V}$ and $\langle u, \alpha \rangle$ an \mathbb{X} -valuation on a ϱ -rational word $u \in \Lambda^{\text{LO}}$. For every infinite sequence $(q_i)_{i \in \mathbb{N}} \in \text{dom}(u)^\mathbb{N}$ there exists a position $q \in \text{dom}(u)$ such that for all $n \in \mathbb{N}$ there are arbitrarily large $i \in \mathbb{N}$ with $\langle u, \alpha[x/q_i] \rangle \lesssim_{\mathcal{F}_n} \langle u, \alpha[x/q] \rangle$.

The key idea here is the following: by Theorem 3.1 there is a winning strategy in the \mathcal{F}_n -game for every $n \in \mathbb{N}$. Obviously, the winning strategy for the \mathcal{F}_k -game is also a winning strategy for the \mathcal{F}_n -game, if $k \geq n$. So, for every $n \in \mathbb{N}$ there is an infinite number of winning strategies. Lemma 3.5 states that these winning strategies can be combined into a single winning strategy for all \mathcal{F}_n -games. This is the main ingredient for showing how the (finite) formulae in \mathcal{F} define a winning strategy in an infinite game; details to this are analogous to the results for two-variable logic which can be found below.

4 Restriction on Two Variables

Let FO^3 be the subset of formulae in FO which only use variables from $\{x, y, z\}$ and let FO^2 be the subset using only variables from $\{x, y\}$. It is well known that for every

sentence in FO there is an equivalent sentence in FO^3 . The study of $\text{FO}^2 = \text{FO}^2[<]$, therefore, arises as a natural limitation. For fragments $\mathcal{F} \subseteq \text{FO}^2[<]$ the order type ϱ can be simplified to $\sigma = \omega + \zeta + \omega^*$. As it will later on be necessary to distinguish the first n and the last n positions of a^σ from the rest, we will say that the union of those positions forms the n -border. Before we can go into detail there is a need for yet another definition. Let p_1 and p_2 be two elements of the same linear ordering (P, \leq_P) ; define

$$\text{ord}(p_1, p_2) = \begin{cases} < & \text{if } p_1 <_P p_2 \\ = & \text{if } p_1 = p_2 \\ > & \text{otherwise, i. e. } p_1 >_P p_2. \end{cases}$$

The simplification from ϱ to σ manifests in the following variation of Theorem 3.4:

Theorem 4.1. *Let $\mathcal{F} \subseteq \text{FO}^2[<]$ be a fragment. For all σ -rational words $u, v \in \Lambda^{\text{LO}}$ the following are equivalent:*

1. $u \models \varphi$ implies $v \models \varphi$ for all sentences $\varphi \in \mathcal{F}$ and
2. Duplicator has a winning strategy in the \mathcal{F} -game on (u, v) .

While the proof for the implication “2. \Rightarrow 1.” is analogous to Theorem 3.4, the proof for the implication “1. \Rightarrow 2.” differs slightly. It heavily relies on the following lemma:

Lemma 4.2. *Let $\mathcal{F} \subseteq \text{FO}^2[<]$ be a fragment and let $\langle u, \alpha \rangle$ and $\langle u, \beta \rangle$ be \mathbb{X} -valuations on $u = v^\sigma$ for a non-empty σ -rational word $v \in \Lambda^{\text{LO}}$ and $\mathbb{X} \subseteq \{x, y\}$. For $\gamma \in \{\alpha, \beta\}$ and $z \in \mathbb{X}$ let $\gamma(z) = (s_z^\gamma, p_z^\gamma)$, and define $\gamma_v : \mathbb{X} \rightarrow \text{dom}(v)$ via $\gamma_v(z) = s_z^\gamma$. Suppose for $n \in \mathbb{N}$ the following three conditions hold:*

1. $\langle v, \alpha_v \rangle \lesssim_{\mathcal{F}_n} \langle v, \beta_v \rangle$.
2. If p_z^γ is in the n -border for $\gamma \in \{\alpha, \beta\}$ and $z \in \mathbb{X}$, then $p_z^\alpha = p_z^\beta$.
3. $\text{ord}(p_x^\alpha, p_y^\alpha) = \text{ord}(p_x^\beta, p_y^\beta)$ if $\mathbb{X} = \{x, y\}$.

Then we have $\langle u, \alpha \rangle \lesssim_{\mathcal{F}_n} \langle u, \beta \rangle$.

Proof. The proof is given by induction on n .

Base case: $n = 0$. As $\text{dom}(u)$ is infinite, $\langle u, \alpha \rangle \lesssim_{\mathcal{F}_0} \langle u, \beta \rangle$ holds if $\langle u, \alpha \rangle \models \varphi$ implies $\langle u, \beta \rangle \models \varphi$ for any literal $\varphi \in \mathcal{F}$. One should note that, obviously, $\langle u, \alpha \rangle$ can only be a model for a literal, if α – and thereby also β – is defined on the variables appearing in the literal. If $\langle u, \alpha \rangle \models x = y$ then $s_x^\alpha = s_y^\alpha$ and $p_x^\alpha = p_y^\alpha$. This implies $s_x^\beta = s_y^\beta$ by 1. and $p_x^\beta = p_y^\beta$ by 3. For $x < y$ and $x \leq y$ one has to distinguish the cases $p_x^\alpha = p_y^\alpha$ and $p_x^\alpha < p_y^\alpha$, but in both cases it is easy to verify that $\langle u, \beta \rangle \models x < y$, or $x \leq y$ respectively, holds. $\langle u, \beta \rangle \models \lambda(x) = a$ follows for $\langle u, \alpha \rangle \models \lambda(x) = a$ directly via 1. and $u(\alpha(x)) = v(s_x^\alpha)$.

Inductive step: $n > 0$. Without loss of generality, let Spoiler choose variable x , quest $q = (s, p) \in \text{dom}(u) = \text{dom}(v) \times (\mathbb{N} \uplus \mathbb{Z} \uplus -\mathbb{N})$ and quantifier $Q \in \{\exists, \forall, \neg\exists, \neg\forall\}$. The proof for $Q \neq \exists$ is either identical or symmetric to the one for $Q = \exists$, as one can easily verify. Therefore, let $Q = \exists$. If $y \notin \mathbb{X}$, Spoiler’s answer can be equal to q and we are done. Thus, it is safe to assume $y \in \mathbb{X}$. By 1. there has to be

$s' \in \text{dom}(v)$ such that $\langle v, \alpha_v[x/s] \rangle \lesssim_{(\mathcal{F}/\exists x)_{n-1}} \langle v, \beta_v[x/s'] \rangle$ holds. Spoiler's answer (s', r) depends on the value of p_y^β . In all cases, r should be chosen in such a way that $\text{ord}(p, p_y^\alpha) = \text{ord}(r, p_y^\beta)$ is maintained and that r being in the $(n-1)$ -border implies $r = p$. If these conditions are satisfied, induction can be applied on $\alpha[x/q]$ and $\beta[x/(s', r)]$ yielding $\langle u, \alpha[x/q] \rangle \lesssim_{(\mathcal{F}/\exists x)_{n-1}} \langle u, \beta[y/(s', r)] \rangle$ which concludes the proof.

If p_y^β is in the n -border choose $r := p$. This maintains the ordering, as $\text{ord}(p, p_y^\alpha) = \text{ord}(r, p_y^\beta)$ by 2. If p_y^β is not in the n -border, but p is in the $(n-1)$ -border, the choice of $r := p$ trivially maintains the ordering, as well. Finally, if neither p_y^β is in the n -border nor p in the $(n-1)$ -border, choose r as the predecessor or successor of p_y^β or equal to p_y^β depending of the ordering of p and p_y^α . By this choice r cannot be in the $(n-1)$ -border. \square

Next, we restate a proposition of [1] in a simplified version. For the sake of brevity, its proof is omitted.

Proposition 4.3. Let $\mathcal{F} \subseteq \text{FO}^2[<]$ be a fragment, and let $\mathbb{X}_1, \mathbb{X}_2 \subseteq \{x, y\}$ be two mutually disjoint sets of variables. Furthermore, for $i = 1, 2$ let $\langle u_i, \alpha_i \rangle, \langle v_i, \beta_i \rangle$ be \mathbb{X}_i -valuations with $\langle u_i, \alpha_i \rangle \lesssim_{\mathcal{F}} \langle v_i, \beta_i \rangle$. Then $\langle u_1 u_2, \alpha_1 \cup \alpha_2 \rangle \lesssim_{\mathcal{F}} \langle v_1 v_2, \beta_1 \cup \beta_2 \rangle$ holds. \square

This allows for the proof of a lemma which is in many aspects analogous to Lemma 3.5.

Lemma 4.4. Let $\mathcal{F} \subseteq \text{FO}^2[<]$ be a fragment and let $\langle u, \alpha \rangle$ be an \mathbb{X} -valuation on a σ -rational word $u \in \Lambda^{\text{LO}}$ with $\mathbb{X} \subseteq \{x, y\}$. For every infinite sequence $(q_i)_{i \in \mathbb{N}} \in \text{dom}(u)^{\mathbb{N}}$ there is a position $q \in \text{dom}(u)$ such that for every $n \in \mathbb{N}$ there is an infinite set $I \subseteq \mathbb{N}$ with $\langle u, \alpha[x/q_i] \rangle \lesssim_{\mathcal{F}_n} \langle u, \alpha[x/q] \rangle$ for every $i \in I$.

Proof. A position q with the above property is called a $\langle u, \alpha \rangle$ -limit point of the sequence $(q_i)_{i \in \mathbb{N}}$. To proof the lemma we have to show the existence of such a $\langle u, \alpha \rangle$ -limit point for any sequence $(q_i)_{i \in \mathbb{N}} \in \text{dom}(u)^{\mathbb{N}}$. Without loss of generality we can assume that $x \notin \mathbb{X}$, because neither $\alpha[x/q_i]$ nor $\alpha[x/q]$ depends on the value of $\alpha(x)$.

The rest of the proof is given by induction on the σ -rational construction of u .

Base case: u is finite. Trivially, we have that $\text{dom}(u)$ is also finite, and that there is a $q \in \text{dom}(u)$ such that $q = q_i$ for infinitely many $i \in \mathbb{N}$. This q is a $\langle u, \alpha \rangle$ -limit point of $(q_i)_{i \in \mathbb{N}}$.

Inductive step 1: $u = v_1 v_2$ with σ -rational words v_1, v_2 . We can split the valuation $\langle u, \alpha \rangle$ into two valuations $\langle v_1, \beta_1 \rangle$ and $\langle v_2, \beta_2 \rangle$ with $\alpha = \beta_1 \cup \beta_2$. For $\ell = 1$ or $\ell = 2$ there is a set $I \subseteq \mathbb{N}$ such that $q_i \in \text{dom}(v_\ell)$ for all $i \in I$. By induction, the subsequence $(q_i)_{i \in I}$ has a $\langle v_\ell, \beta_\ell \rangle$ -limit point, and, by Proposition 4.3 this limit point is also a $\langle u, \alpha \rangle$ -limit point of $(q_i)_{i \in \mathbb{N}}$.

Inductive step 2: $u = v^\sigma$ for a σ -rational word v . As $\text{dom}(u) = \text{dom}(v) \times (\mathbb{N} \uplus \mathbb{Z} \uplus -\mathbb{N})$, let $q_i =: (s_i, p_i)$ for $i \in \mathbb{N}$. Define $\alpha_v : \mathbb{X} \rightarrow \text{dom}(v)$ as the restriction of α to values in $\text{dom}(v)$ by discarding the $(\mathbb{N} \uplus \mathbb{Z} \uplus -\mathbb{N})$ -part. In each of the following cases there is a definition of a value p and an infinite index set J such that induction yields a

$\langle v, \alpha_v \rangle$ -limit point s of the sequence $(s_i)_{i \in J}$. Define $I \subseteq J$ as the infinite set of indices i with $\langle v, \alpha_v[x/s_i] \rangle \lesssim_{\mathcal{F}_n} \langle v, \alpha_v[x/s] \rangle$ and define q as (s, p) . One should observe that in all cases p is chosen in a way that satisfies the conditions of Lemma 4.2. Thus, $\langle u, \alpha[x/(s_i, p_i)] \rangle \lesssim_{\mathcal{F}_n} \langle u, \alpha[x/(s, p)] \rangle$ follows for all $i \in I$.

Case 1: There is p such that $p = p_i$ for infinitely many i . Let J be the infinite set of these i .

Case 2: There is an unbounded subsequence of $(p_i)_{i \in \mathbb{N}}$ with values in \mathbb{N} or one with values in $-\mathbb{N}$. The argumentation for values in $-\mathbb{N}$ is dual to the one with values in \mathbb{N} . Therefore, demonstration is only given for the \mathbb{N} -case. If $\alpha(y) = (s_y, p_y)$ is defined and if p_y is in the \mathbb{N} -part, let m be the maximum of n and p_y , and let $p = 0$ in the \mathbb{Z} -part. If p_y is in the \mathbb{Z} -part, let $m = n$ and $p = p_y - 1$. Otherwise, let $m = n$ and $p = 0$ in the \mathbb{Z} -part. Let J be the infinite set of those indices i of the unbounded subsequence with $p_i > m$. Special note shall be given to the fact that, by choice, the ordering of p_i and p_y is equal to the one of p and p_y for all $i \in J$.

Case 3: There is an unbounded subsequence of $(p_i)_{i \in \mathbb{N}}$ with values in \mathbb{Z} . By duality one can assume that the subsequence is upwards unbounded. If $\alpha(y) = (s_y, p_y)$ is defined and p_y is in the \mathbb{Z} -part, let $p = p_y + 1$. Otherwise, choose p as the 0 of the \mathbb{Z} -part. Finally, choose J as the infinite set of all indices i in the upwards unbounded subsequence with $p_i \geq p$. Again, this choice ensures that the ordering of p_i and p_y is equal to the one of p and p_y for all $i \in J$. \square

One of the major differences between Lemma 4.4 and Lemma 3.5 is that in the latter case one can take a more local view since if we factorize $\varrho = \varrho_1 + 1 + \varrho_2$ with infinite factors ϱ_1 and ϱ_2 , then both ϱ_1 and ϱ_2 are isomorphic to ϱ . We cannot apply this simplification here in Lemma 4.4.

With this lemma in place, proving the implication “1. \Rightarrow 2.” of Theorem 4.1 is a simple matter of adapting the corresponding parts of the proof for Theorem 3.4.

Proof of Theorem 4.1, “1. \Rightarrow 2.”. We call a configuration $(\mathcal{G}, \langle u, \alpha \rangle, \langle v, \beta \rangle)$ of the \mathcal{F} -game on (u, v) good for Duplicator if $\langle u, \alpha \rangle \lesssim_{\mathcal{G}_n} \langle v, \beta \rangle$ for every $n \in \mathbb{N}$. Good configurations especially satisfy $\langle u, \alpha \rangle \lesssim_{\mathcal{G}_0} \langle v, \beta \rangle$, so Spoiler does not win immediately in good configurations. Thus, staying in good configurations forms a winning strategy for Duplicator. By Theorem 3.1 the initial configuration (\mathcal{F}, u, v) is good.

Now suppose we are in the good configuration $(\mathcal{G}, \langle u, \alpha \rangle, \langle v, \beta \rangle)$ and Spoiler chooses quantifier Q , variable x and quest q . Now, Duplicator has to choose such an answer that the follow-up configuration is good again. This choice, obviously, depends on Q ; here, we only demonstrate the case for $Q = \exists$, as the other cases are similar. For $Q = \exists$ Spoiler’s quest q has to be in $\text{dom}(u)$. Because we are in a good configuration, we have $\langle u, \alpha \rangle \lesssim_{\mathcal{G}_{n+1}} \langle v, \beta \rangle$ for every $n \in \mathbb{N}$. This means that for every $n \in \mathbb{N}$ Duplicator has an answer $r_n \in \text{dom}(v)$ such that $\langle u, \alpha[x/q] \rangle \lesssim_{\mathcal{G}_{n+1}/\exists x} \langle v, \beta[x/r_n] \rangle$ holds. Because of $\mathcal{G}_{n+1}/\exists x = (\mathcal{G}/\exists x)_n$ this is equivalent to $\langle u, \alpha[x/q] \rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r_n] \rangle$. Lemma 4.4 yields an $r \in \text{dom}(v)$ such that for every $n \in \mathbb{N}$ there is an infinite set $I \subseteq \mathbb{N}$ with $\langle v, \alpha[x/r_i] \rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \alpha[x/r] \rangle$ for all $i \in I$. Choosing this r

as Duplicator's answer results in a good configuration: for every $n \in \mathbb{N}$ there is an $i \geq n$ with $\langle v, \beta[x/r_i] \rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r] \rangle$ and $\langle u, \alpha[x/q] \rangle \lesssim_{(\mathcal{G}/\exists x)_i} \langle v, \beta[x/r_i] \rangle$ holds by choice of r_i . This implies $\langle u, \alpha[x/q] \rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r] \rangle$ and, in turn, $\langle u, \alpha[x/q] \rangle \lesssim_{(\mathcal{G}/\exists x)_n} \langle v, \beta[x/r] \rangle$. \square

Above FO^2 and below σ . While it is sufficient to consider σ -rational words for fragments $\mathcal{F} \subseteq \text{FO}^2[<]$, this restriction in consideration is insufficient for FO^3 and, hence, for FO . To see this, consider $u = a^\sigma$ and $v = a^\sigma a^\sigma = a^{\omega+\zeta+\zeta+\omega^*}$ for an $a \in \Lambda$. These two words cannot be distinguished by any formula in FO , since there is a winning strategy for Duplicator in the FO_n -game for every $n \in \mathbb{N}$. Therefore, they especially cannot be distinguished by any formula in FO^3 . On the other hand, there is a winning strategy for Spoiler in the infinite FO^3 -game on (u, v) : In the first round, Spoiler can choose variable x , quantifier \forall and a quest in the first \mathbb{Z} -part of $\text{dom}(v)$. Duplicator's answer has to be in the \mathbb{Z} -part of $\text{dom}(u)$, since otherwise there would be only finitely many positions left on one side which would allow Spoiler to win in finitely many rounds. In the second round, Spoiler can choose variable y , quantifier \forall and a quest in the second \mathbb{Z} -part of $\text{dom}(v)$. Again, Duplicator's answer needs to be in the \mathbb{Z} -part of $\text{dom}(u)$ for the same reason as before. But: this leaves only finitely many positions between Duplicator's first and second answer in $\text{dom}(u)$, while there still are infinitely many positions between Spoiler's two quests. Therefore, Spoiler can win in finitely many rounds.

On the other hand, one might assume that it is possible to further restrict σ to $\sigma' = \omega + \omega^*$, the order type of $a^\omega a^{\omega^*}$ for $a \in \Lambda$, in the case of fragments $\mathcal{F} \subseteq \text{FO}^2[<]$. But this is not the case, as can be seen when considering the words $u = a^{\sigma'}$ and $v = a^{\sigma'} a^{\sigma'} = a^{\omega+\zeta+\omega^*}$. In $\text{dom}(v)$ there are positions which are infinitely far away from both ends of the word, namely the positions belonging to the \mathbb{Z} -part, but there are no such positions in $\text{dom}(u)$. This constitutes a winning strategy for Spoiler in the FO^2 -game on (u, v) .

Applications. Let $\mathcal{F} \subseteq \text{FO}^2[<]$ be a fragment and let $s, t \in T_\Lambda$. One of the main applications of Theorem 4.1 is the equivalence of the following two properties:

1. The identity $s = t$ holds in the syntactic monoid of every language definable in \mathcal{F} .
2. Duplicator has winning strategies in the \mathcal{F} -games on $(\llbracket s \rrbracket_\sigma, \llbracket t \rrbracket_\sigma)$ and $(\llbracket t \rrbracket_\sigma, \llbracket s \rrbracket_\sigma)$.

Here, $\llbracket s \rrbracket_\sigma$ is the generalized word obtained by replacing π by σ . The proof is identical to the corresponding statement in [1]; a definition of the syntactic monoid can also be found in this paper.

References

- [1] M. Huschenbett and M. Kufleitner. Ehrenfeucht-Fraïssé games on omega-terms. In *STACS 2014, Proceedings*, volume 25 of *LIPIcs*, pages 374–385. Dagstuhl Publishing, 2014.
- [2] M. Kufleitner and A. Lauser. Lattices of logical fragments over words. In *ICALP 2012, Proceedings Part II*, volume 7392 of *LNCS*, pages 275–286. Springer, 2012.
- [3] D. Perrin and J.-É. Pin. *Infinite words*, volume 141 of *Pure and Applied Mathematics*. Elsevier, 2004.